ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Technology 6 (2012) 98 – 107

Procedia Technology

2nd International Conference on Communication, Computing & Security [ICCCS-2012]

FPGA Based Real Time Human Hand Gesture Recognition System

Amit Gupta[a], Vijay Kumar Sehrawat[b], Mamta Khosla[c], a[*]

[a,b,c]Department of Electronics and Communication Engineering, Dr. B.R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

Abstract

This work proposes a real time human hand gesture recognition system for human computer interaction. The proposed system can recognize 10 different hand gestures at faster rate with reasonable accuracy. The gestures are classified on the basis of shape-based features. Four different shape based features are used for better accuracy. The illumination compensation technique is employed for robust recognition under varying background lightning conditions. Skin color segmentation is used to minimize the chances of false detection. The proposed system is modeled using Verilog HDL and targeted for Xillinx Virtex2 Pro FPGA board. The accuracy of the system is computed as 94.40%.

Keywords: Hand Gesture Recognition; Human Computer Interface; FPGA.

1. Introduction

Human hand gestures can be used as an important communication tool for human computer interaction [7]. Human hand gestures integrated with vocal language and facial expressions make the communication process more interactive and convenient for user. There are many application areas for human hand gestures based human computer interface including virtual gaming, security, sign language recognition [3][13][14].

Human Computer Interaction using human hand gestures requires a human hand recognition system that can recognize and classify variety of the hand gestures in real time. But a lot of computations are required to

* Corresponding author. Tel.: +919417033066
E-mail address: amit.gupta326@gmail.com

doi:10.1016/j.protcy.2012.10.013

recognize the hand gestures as proposed by various researchers [10] [11] [12]. The techniques proposed in [10] [12] require a huge training data which in turn increases the processing time and hence, limits their use for real time applications.

The proposed hand gesture recognition system addresses the above stated problems for real time recognition of human hand gestures. The proposed system is based on the computation of simple shape based features to recognize the hand gestures. Further the FPGA implementation of the system requires very less processing time and qualifies it for use in the real time human computer interaction. The proposed system employs illumination compensation for the varying background lightning conditions for better accuracy. Also the skin color segmentation is used for segmentation of hand to reduce the chances of false segmentation of other objects in the background.

The paper is organized as follows. In section 2, related works of different authors are discussed. Section 3 describes the proposed scheme for the hand gesture recognition system. Section 4 describes the FPGA Implementation. Finally results are shown and conclusion is drawn in section 5.

2. Related works

Various hand gesture recognition algorithms have been developed by researchers for different application areas including human computer interaction, virtual game control, robotics and others [3][10][14].

Ishikawa [1] developed a system for recognition of hand gestures based on self-organization technique, but the system requires the user to wear a data glove for hand segmentation, thereby limiting the freedom of user to communicate with bare hands. Erdem [2] proposed a method to recognize the hand based on the shape, the work can be further enhanced by calculating some more shape-based features to recognize the different hand gestures. Gupta [3] presented an algorithm to recognize hand gestures using contour analysis of hand and requires a lot of mathematical computations, the system can be further optimized for real time processing. Amornched [11] proposed a shape-based scheme for hand gesture recognition but they used fixed threshold for hand segmentation which is susceptible to false segmentation of other bright objects. Tsui. [13] proposed an FPGA based smart camera for gesture analysis. However, the performance of their system can be further improved by high resolution images and adding more features.

The proposed human hand gesture recognition system addresses the limitations of the earlier methods. The proposed system is based on the simple shape based features of the hand gestures. The system requires the computation of 4 different features for better accuracy and FPGA implementation ensures less processing time for real time applications.

3. Proposed hand gesture recognition system

The proposed gesture recognition system is designed to recognize 10 different static hand gestures as shown in Fig. 1. The gestures are classified into two types: with thumb in Fig. 1(a)-(d) and without thumb in Fig. 1(e)-(j). Further the proposed system is able to recognize the hand gestures irrespective of the orientation of hand.

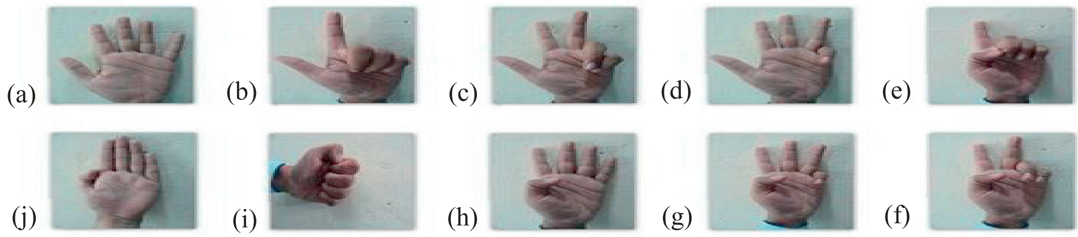

Fig. 1. (a) –(j) hand gestures defined for the proposed system, from top left going clockwise (a) – (j).

Fig. 2 shows the scheme of proposed hand gesture recognition system. The system can be divided into three parts: Image Acquisition, Pre-Processing and Gesture Recognition.

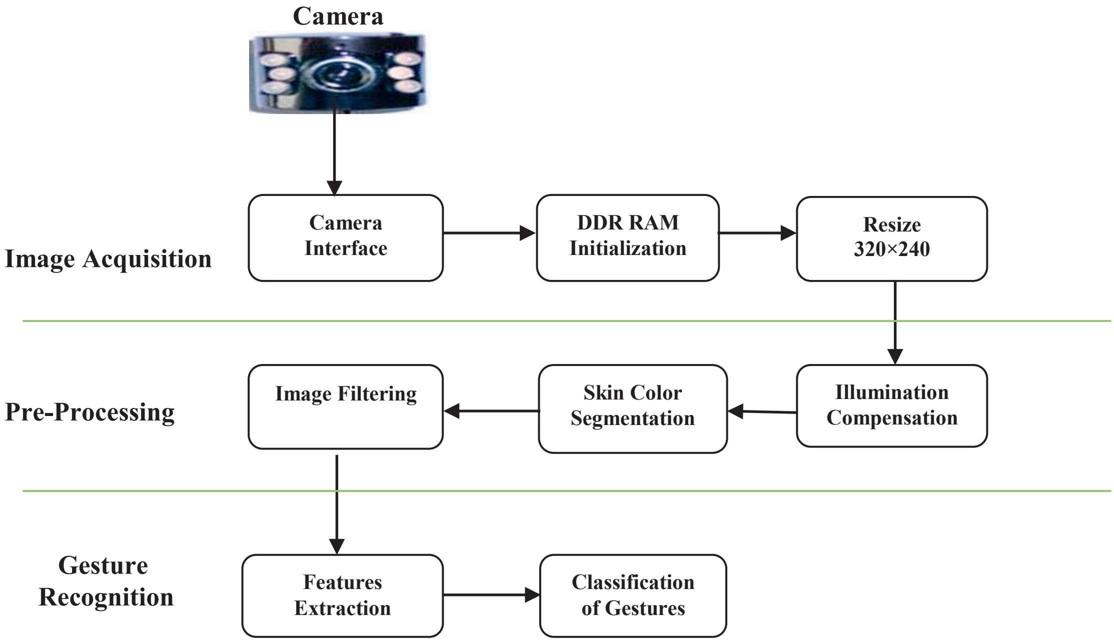

Fig. 2. proposed hand gesture recognition system.

3.1. Image acquisition

The image of hand for processing is acquired using CMOS image sensor camera, which generates the *R,G,B* components for the image. The mathematical model of image acquisition process [5][6] is represented in (1)-(3).

$$R = \sigma \int E(\lambda)S(\lambda)Q_R(\lambda)d\lambda \tag{1}$$

$$G = \sigma \int E(\lambda)S(\lambda)Q_G(\lambda)d\lambda \tag{2}$$

$$B = \sigma \int E(\lambda)S(\lambda)Q_B(\lambda)d\lambda \tag{3}$$

Where, σ denotes a constant factor, $E(\lambda)$ is spectral distribution of light, $S(\lambda)$ is the surface reflectance of the hand and other objects in the region and $Q_R(\lambda)$, $Q_B(\lambda)$, $Q_G(\lambda)$ are spectral sensitivities of the CMOS image sensor.

The captured image is retrieved by the camera interfacing module in *R,G,B* components. The camera interfacing module is designed to setup communication between camera and the FPGA board. The image is then stored on board DDR RAM and magnified to the 320×240 pixels resolution around the region of interest (ROI) as the resolution of the camera may be set to different value. The resolution of the camera may be adjusted manually. The distance between the camera and hand is about 35 centimeters and the hand gesture is held in predefined region of interest to avoid invariance due to location of hand w.r.t. camera. The FPGA is operated at 100 MHz, which allows acquiring the images in real time.

3.2. Pre-processing and segmentation

The quality of the image acquired depends upon background lightning conditions and quality of imaging device. A low quality image affects the accuracy of the system. So, pre-processing of the image is mandatory to enhance the quality of the acquired image, which ultimately improves the accuracy of the proposed system. The high quality enhanced image is then processed to segment the hand information from the image. The image is then filtered for removal the noisy patches to obtain the good quality image.

In the proposed hand gesture recognition system, illumination compensation of the *R,G,B* components of the acquired image has been done to compensate the effects of varying background lightning conditions, skin color segmentation is employed for hand segmentation and morphological image filtering [4] is used to filter the image.

3.2.1. Illumination compensation

The spectral distribution, *E(λ)* as described in (1) of light source changes with change in the environment conditions like indoor, outdoor, shadows, direct sunlight which degrades the quality of hand's image acquired by the imaging device. These illumination changes are compensated to get good quality images. The mathematical description of the illumination compensation is as follows:

$$R(Mean) = \frac{1}{320\times240}\sum_{X=1}^{240}\sum_{Y=1}^{320} I(X,Y,1) \tag{4}$$

$$G(Mean) = \frac{1}{320\times240}\sum_{X=1}^{240}\sum_{Y=1}^{320} I(X,Y,2) \tag{5}$$

$$B(Mean) = \frac{1}{320\times240}\sum_{X=1}^{240}\sum_{Y=1}^{320} I(X,Y,3) \tag{6}$$

While, *I(x,y,z)* is the 3D matrix representing the acquired image where, x denotes the row index, y denotes the column index and z represents the color component *R,G,B* as obtained in (1). First, the mean of the *R,G,B* values is computed in accordance with (4)-(6).

The minimum among the three parameters computed in (4),(5) and (6) is denoted as ε and is defined as:

$$\varepsilon = \text{minimum}\ (\text{R(Mean)}, \text{G(Mean)}, \text{B(Mean)}) \tag{7}$$

The R,G,B components after illumination compensation is calculated as in (8),(9) and (10), which brings about the uniformity in the image.

$$R = \frac{R\times R(Mean)}{\varepsilon} \tag{8}$$

$$G = \frac{G\times G(Mean)}{\varepsilon} \tag{9}$$

$$B = \frac{B\times B(Mean)}{\varepsilon} \tag{10}$$

3.2.2. Skin color segmentation

Segmentation is the process of partitioning the digital image into multiple segments [7]. Segmentation of image enables the system to retain the useful parts of the image while discarding the other parts. Pixels representing the hand gesture is the useful information for the processing in hand gesture recognition system. So, the segmentation module is designed to extract the hand from the whole image captured by the camera.

The scheme of segmentation used in the system is based on the skin color model obtained by analysing the skin color pixels of the hand in YCbCr color space as shown in Fig. 3. Here, the shaded region shows the expected skin color pixels in the image.

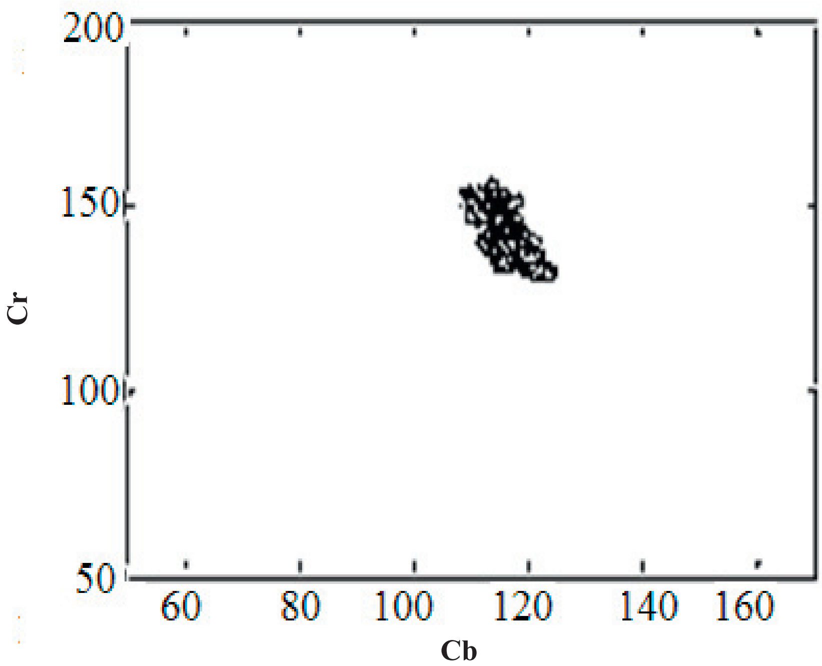

Fig. 3. skin color model in YCbCr color space.

Skin color segmentation scheme gives better results as compared to the Otsu method [9] which uses a fixed threshold on the basis of gray level histogram. We compared both segmentation techniques and the comparative segmentation results for Otsu method and the skin color segmentation are shown in Fig. 4. Fig. 4(a) shows the original hand images, Fig. 4(b),4(c) shows the segmentation using Otsu method and skin segmentation respectively.

3.2.3 Image filtering

Skin color segmentation can effectively segment the skin color pixels, but it gives noisy holes in the image due to nails and other hand accessories like ring. So, morphological image filtering is used to filter

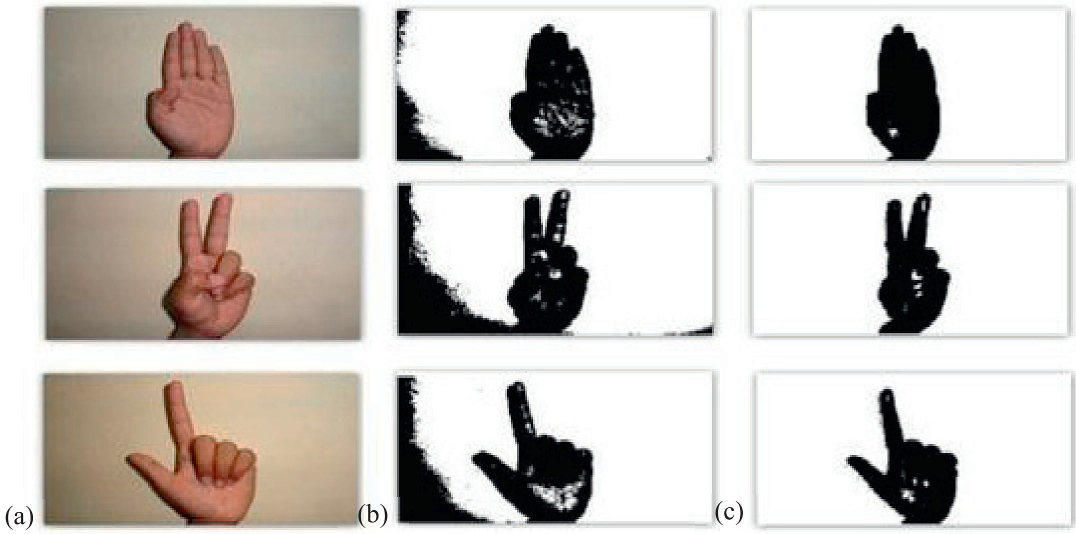

Fig. 4. comparative segmentation for Otsu method and skin color segmentation, (a) original images, (b) results for Otsu method, (c) results for skin color segmentation

out the noisy pixels. The morphological operation performed on the image is described in (11)

$$A \circ B = (A \ominus B) \oplus B \tag{11}$$

Here, A represents the image being filtered, B is the structuring element as shown in Fig. 5

0	1	0
1	1	1
0	1	0

Fig. 5. structuring element B used in Eq. (11).

Fig. 6 shows the results obtained by applying the image filtering techniques. The filtered images are free from any noisy patches. Fig. 6(a) and 6(b) show the noisy and the filtered images respectively.

3.3. Gesture recognition

Human hands can show different gestures depending upon the shape of the hand pattern. The different hand gestures can be distinguished on the basis of shape based features. In this section, four different shape based features viz. area, perimeter, thumb detection and radial profile of hand are described; and the proposed criterion for classification for hand gestures is defined.

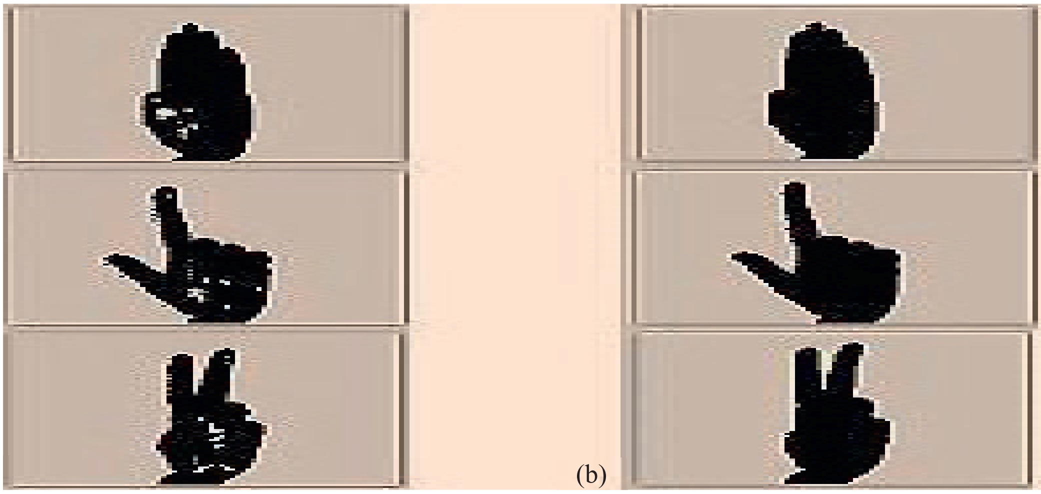

Figure. 6. results of morphological filtering, (a) noisy Images after skin color segmentation, (b) filtered images after morphological filtering.

3.3.1 Features extraction

Shape based recognition requires very less computation effort and saves a lot of FPGA area as compared to other approaches [10]. But the approach is not very popular as a hand can assume many shapes. So, in the proposed hand gesture recognition system four different shape based features are calculated rather than relying on a single feature. The different shape based features used in the proposed system are described as follows:

3.3.1.1 Area of hand

Hand can occupy different areas for different hand getures. So, the different gestures can be differentiated on the basis of area. The area of hand gestures for 10 different individuals was analysed and the results are shown in Fig. 7(a). The areas of the hand gestures are computed in accordance with (12).

$$Area\ of\ Hand = \sum_{x=0}^{240} \sum_{y=0}^{320} I(x,y)(segmented) \tag{12}$$

3.3.1.2 Perimeter of hand

Perimeter of the image can be used to distinguish the hand gestures as Fig. 1(a) and Fig. 1(j) have the same areas but can be distinguished on the basis of perimeter. Eq. (13) describes the perimeter of the image. Fig. 7(b) shows the perimeter values for the 10 different hand gestures.

$$Perimeter = \sum_{x=0}^{240} \sum_{y=0}^{320} Edge(I(x,y)) \tag{13}$$

3.3.1.3 Thumb detection

The peninsular shape of thumb adds additional area to sides of the hand, so the presence of the thumb is detected by computing the area of the hand enclosed in the thumb detection boxes and comparing it with the area of whole hand. Fig. 8 shows the thumb detection boxes to calculate extra area due to presence of thumb.

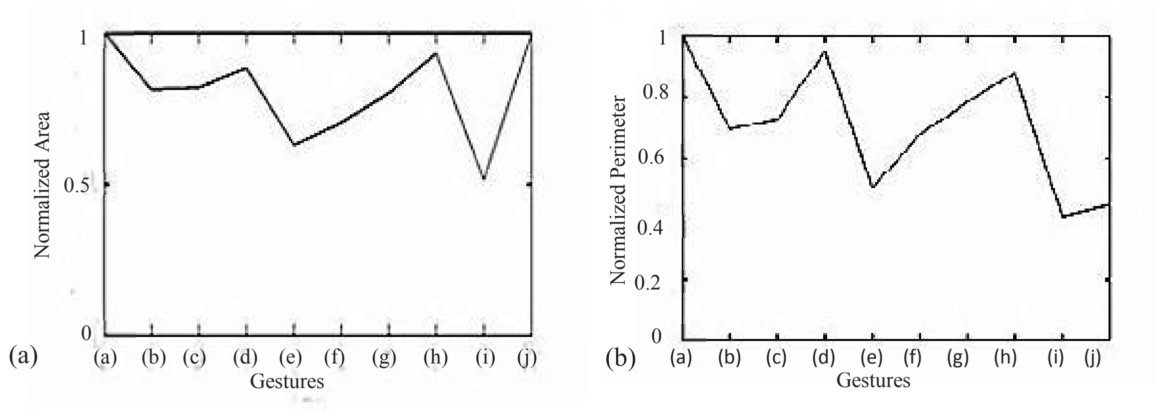

Fig. 7. (a) variation of area for different hand gestures; (b) variation of perimeter for different hand gestures.

We analysed a set of images of hands of 25 different people. The results of our analysis shows that presence of thumb adds 6.90% of the total area of hand.

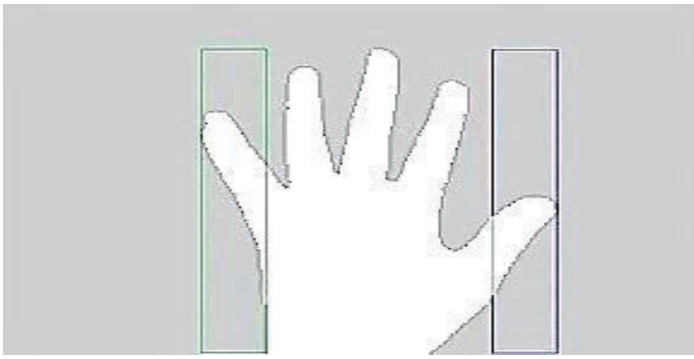

Fig. 8 thumb detection boxes

3.3.1.3 Radial profile and angular position

The radial profile of hand is computed by plotting the Euclidean distance of hand edges from the reference point within the hand. Eq.(14) gives the mathematical description of Euclidean distance.

$$E.D.(p,q) = \sqrt{(x-p)^2 + (y-q)^2} \tag{14}$$

Here, *(x,y)* is the refrence point within hand and *(p,q)* are edge pixels of hand. The plot of radial profile gives the unique peaks for the fingers. Number of fingers in the hand gesture are equal to the number of peaks above a threshold value for Euclidean Distance. The threshold value for the peaks detection is set to be 75% of the maximum value of Euclidean distance. Also, angular position of the peaks w.r.t. wrist recognizes the raised fingers. Fig. 9 shows the radial profiles for the hand gestures with peaks at certain angles.

3.3.2 Classification of gestures

The hand gestures shown in Fig. 1 are classified according to the classification criterion defined in Table 1. The classification criterion is defined by analysing the images of hands of 25 individuals. The proposed criterion is tested for different people and the Matlab implementation of the proposed algorithm demonstrates an accuracy of 94.40%.

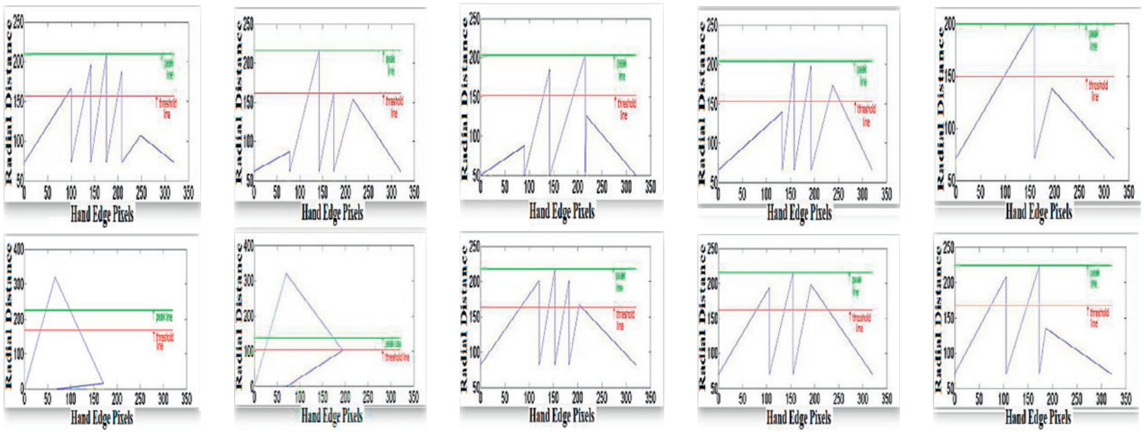

Fig. 9. radial distance profiles with angular position for hand gestures shown in Fig. 1 (a)-(j), strating from top left going clock wise.

Table 1. Proposed criterion for classification of hand gestures

Hand Gesture	Thumb	Normalized Area	Normalized Perimeter	Peaks
Fig. 1(a)	Yes	>0.9	>0.8	4
Fig. 1(b)	Yes	0.5-0.8	0.61-0.83	1
Fig. 1(c)	Yes	0.5-0.8	0.61-0.83	2
Fig. 1(d)	Yes	0.5-0.8	>0.8	3
Fig. 1(e)	No	0.5-0.8	0.61-0.83	1
Fig. 1(f)	No	0.5-0.8	0.61-0.83	2
Fig. 1(g)	No	0.5-0.8	0.61-0.83	3
Fig. 1(h)	No	>0.9	>0.8	4
Fig. 1(i)	No	<0.5	<0.5	1
Fig. 1(j)	No	>0.9	<0.5	1

4. FPGA implementation

The proposed hand gesture recognition system is modelled in Verilog HDL and synthesized using Xillinx® ISE. The target FPGA board used is the Virtex2 Pro FPGA development board. Hand gesture recognition system needs to store the image in the form of huge data arrays. The on board DDR RAM of Virtex2 Pro board is used for the storage of captured image. The FPGA is operated on the clock frequency of 100 MHz. Table 2. shows the results of FPGA implementation of proposed hand gesture recognition system.

Table 2. Results of FPGA implementation

Parameter	Value
FPGA	Virtex2 Pro
Image Resolution	240*320
Area	4.7%
Recognition Time	2.1msec
Clock Frequency	100 MHz

5. Results and conclusion

The proposed system is tested for a real time database comprising images of hands of 25 different people in different age groups to check its reliability and robustness. The accuracy of the proposed system for different hand gestures is given in Table. 3 that shows that accuracy of the proposed system is relatively less for the gestures shown in Fig. 1(i) and 1(j) which limits the overall accuracy of the system to the 94.40 %.

The accuracy of the proposed system is fairly better as compared to the other hand gesture recognition systems given by [11][12]. Also, the goal of the proposed system is the real time recognition of hand gestures. The recognition time for the system is 2.1 msec and is suitable for its use in the real time systems.

Table 3. Accuracy of the proposed system

Hand Gesture	Total Images	Recognised Correctly	Accuracy (%)
Fig. 1(a)	25	25	100
Fig. 1(b)	25	25	100
Fig. 1(c)	25	25	100
Fig. 1(d)	25	25	100
Fig. 1(e)	25	23	92
Fig. 1(f)	25	25	100
Fig. 1(g)	25	24	96
Fig. 1(h)	25	25	100
Fig. 1(i)	25	18	72
Fig. 1(j)	25	21	84
Total	250	236	94.40

References

[1] Ishikawa, M.; Matsumura, H., "Recognition of a hand-gesture based on self-organization using a DataGlove ", *Proc. Of 6th International Conference on Neural Information Processing,* 1999.
[2] Erdem Yrk, E. Konukoglu, B. Sankur , and J. Darbon, "Shape-Based Hand Recognition," *IEEE Transaction On Image Processing*, Vol. 15, No. 7, page 1803-1815, July 2006.
[3] Gupta, L.; Suwei Ma; , "Gesture-based interaction and communication: automated classification of hand gesture contours," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol.31, no.1, pp.114-120, Feb 2001.
[4] Jain, A.K. ,"Fundamentals of Digital Image Processing", Prentice Hall 2010.
[5] Finlayson, G.D.; Hordley, S.D.; Cheng Lu; Drew, M.S.; , "On the removal of shadows from images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.28, no.1, pp.59-68 Jan 2006.
[6] Ukil Yang; Bongjoe Kim; Kwanghoon Sohn; , "Illumination invariant skin color segmentation," *4th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2009.* vol., no., pp.636-641, 25-27 May 2009.
[7] www.wikipedia.com.
[8] Son Lam Phung; Bouzerdoum, A.; Chai, D.; , "A novel skin color model in YCbCr color space and its application to human face detection," *International Conference on Image Processing. 2002. Proceedings. 2002*, vol.1, no., pp. I-289- I-292 vol.1, 2002.
[9] Otsu N,A thresholding selection method from gray-level histograms IEEE Trans System Man Cybernetic.1979, pp.62-66.
[10] Chang Tan; Xiao, Nanfeng; , "Improved RCE neural network and its application in human-robot interaction based on hand gesture recognition," *2nd International Conference on Information Science and Engineering (ICISE), 2010*, vol., no., pp.1260-1263, 4-6 Dec. 2010.
[11] Jinda-apiraksa, Amornched; Pongstiensak, Warong; Kondo, Toshiaki; , "A simple shape-based approach to hand gesture recognition," *International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 2010*, vol., no., pp.851-855, 19-21 May 2010.
[12] Chen-Chiung Hsieh; Dung-Hua Liou; Lee, D.; , "A real time hand gesture recognition system using motion history image," *2nd International Conference on Signal Processing Systems (ICSPS), 2010*, vol.2, no., pp.V2-394-V2-398, 5-7 July 2010.
[13] Tsui, T.; Yu Shi; , "An FPGA-based Smart Camera for Gesture Analysis for Healthcare Applications," *International Conference on Consumer Electronics,(ICCE), 2008. Digest of Technical Paper.*, vol., no., pp.1-2, 9-13 Jan. 2008.
[14] Rautaray, S.S.; Agrawal, A.; , "Interaction with virtual game through hand gesture recognition," *International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT), 2011*, vol., no., pp.244-247, 17-19 Dec. 2011